



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### Survey of type 6 group variants of hepatitis C virus in Southeast Asia by using a core-based genotyping assay

**Citation for published version:**

Mellor, J, Walsh, EA, Prescott, LE, Jarvis, LM, Davidson, F, Yap, PL & Simmonds, P 1996, 'Survey of type 6 group variants of hepatitis C virus in Southeast Asia by using a core-based genotyping assay', *Journal of Clinical Microbiology*, vol. 34, no. 2, pp. 417-23.

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher's PDF, also known as Version of record

**Published In:**

Journal of Clinical Microbiology

**Publisher Rights Statement:**

Copyright 1996, American Society for Microbiology

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



## Survey of Type 6 Group Variants of Hepatitis C Virus in Southeast Asia by Using a Core-Based Genotyping Assay

J. MELLOR,<sup>1</sup> E. A. WALSH,<sup>2</sup> L. E. PRESCOTT,<sup>2</sup> L. M. JARVIS,<sup>1</sup> F. DAVIDSON,<sup>2</sup> P. L. YAP,<sup>2</sup>  
P. SIMMONDS,<sup>1\*</sup> AND INTERNATIONAL HCV COLLABORATIVE STUDY GROUP†

*Department of Medical Microbiology, University of Edinburgh, Edinburgh EH8 9AG,<sup>1</sup> and Edinburgh and South East Scotland Blood Transfusion Service, Royal Infirmary of Edinburgh, Edinburgh EH3 9HB,<sup>2</sup> Scotland*

Received 11 August 1995/Returned for modification 13 October 1995/Accepted 3 November 1995

Previous surveys of the prevalences of genotypes of hepatitis C virus (HCV) in different populations have often used genotyping assays based upon analysis of amplified sequences from the 5' noncoding region (5'NCR), such as restriction fragment length polymorphism (RFLP) or hybridization with type-specific probes (e.g., InnoLipa). Although highly conserved, this region contains several type-specific nucleotide polymorphisms that allow major genotypes 1 to 6 to be reliably identified. Recently, however, novel HCV variants found in Vietnam and Thailand that are distantly related to the type 6a genotype (type 6 group) by phylogenetic analysis of coding regions of the genome often have sequences in the 5'NCR that are similar or identical to those of type 1 and could therefore not be identified by any assay of sequences in this region. We developed a new genotyping assay based upon RFLP of sequences amplified from the more variable core region to investigate their distribution elsewhere in southeast (SE) Asia. Among 108 samples from blood donors in seven areas that were identified as type 1 by RFLP in the 5'NCR, type 6 group variants were found in Thailand (7 from 28 samples originally identified as type 1) and Burma (Myanmar) (1 of 3) but were not found in Hong Kong ( $n = 43$ ), Macau ( $n = 8$ ), Taiwan ( $n = 6$ ), Singapore ( $n = 2$ ), or Malaysia ( $n = 18$ ). Although this small survey suggests a relatively limited distribution for type 6 group variants in SE Asia, larger studies will be required to explore their distribution in other geographical regions and the extent to which their presence would limit the practical usefulness of 5'NCR-based genotyping assays for clinical or epidemiological purposes.

Hepatitis C virus (HCV) is a positive-stranded RNA virus of approximately 9,400 nucleotides and has been shown to be the major etiologic agent of parentally transmitted non-A, non-B hepatitis (7). Sequence comparisons of variants from different geographical areas have led to the identification and classification of at least six major genotypes, many of which contain a number of more closely related, yet distinct subtypes of the virus (5, 20, 24). Phylogenetic analysis of complete genomic sequences (16) or even relatively short subgenomic regions (core [6, 15, 24], E1 [5], NS-4 [2, 21], or NS-5 [11]) may be used for virus classification into genotypes. The overall sequence similarities over complete genomic sequences are at least 91% within variants of the same genotype, approximately 79% (range, 77 to 80%) between subtypes, and about 68% (range, 66 to 69%) between different types (26).

Infection with different genotypes may produce clinically relevant differences in the liver disease caused by HCV. For example, there are now several reports describing a greater

sensitivity of HCV genotypes 2 and 3 to interferon treatment than is found in type 1-infected patients (for recent reviews, see references 4 and 19). Assays that can identify the infecting genotype without having to sequence the virus directly are therefore becoming important in patient management and in epidemiological studies of HCV transmission. Published genotyping methods fall into one of two categories, those based upon direct virus detection by reverse transcription-PCR, followed by analysis of the amplified DNA, or indirect (serological) assays, in which the infecting genotype is inferred from the pattern of antibody reactivity to type-specific epitopes in the core or NS-4 region (2, 14, 21, 25).

Although the 5' noncoding region (5'NCR) of HCV is highly conserved between genotypes, it is an attractive target for genotyping assays, because of the strong association of specific nucleotide polymorphisms with genotype (for a recent review, see reference 22). These genotypes may be readily identified by restriction endonucleases or by type-specific probes (9, 27), although the accuracy for the detection of genotypes 1 to 6 is limited to 96 and 84% for the two assays, respectively (22), mainly because of misidentification of more recently described genotypes.

Recently, several new variants have been identified among blood donors and hepatitis C patients in Vietnam (26) and Thailand (1, 15). Sequence comparisons of the core, NS-5, and E1 regions indicated that these new genotypes often showed sufficient nucleotide differences from existing variants to be classified as new major genotypes. However, we have found that these novel variants group with type 6 upon phylogenetic analysis in both the E1 and NS-5 regions (15), producing what may equally well be interpreted as a single, highly diverse single genotype that includes the previously described type 6a

\* Corresponding author. Mailing address: Department of Medical Microbiology, University of Edinburgh, Teviot Place, Edinburgh EH8 9AG, Scotland. Phone: (44) 131 650 3138. Fax: (44) 131 650 6531.

† M. J. Nowicki and J. W. Mosley, Transfusion Safety Study, University of Southern California School of Medicine, Los Angeles, Calif.; C. K. Lin and C. L. Lai, Red Cross Blood Transfusion Service, Yau-matei, Hong Kong; G. de Olim and I. A. Martins, Macau Blood Transfusion Centre, Macau; G. Duraisamy, National Blood Services Centre, Hospital Kuala Lumpur, 50586 Kuala Lumpur, Malaysia; Y. W. Ong and D. Teo, Singapore Blood Transfusion Service, National Blood Centre, Singapore 0316, Singapore; M. Lin, Transfusion Medicine Laboratory, National Health Research Institute, Mackay Memorial Hospital, Taipei, Taiwan; and C. Nuchprayoon and S. Tanprasert, Thai Red Cross Society, Bangkok 10330, Thailand.

genotype found in Hong Kong and Macau (5, 9, 20). These new variants are referred to as type 6 group variants in this study.

Although type 6a and all of the previously described major genotypes (types 2 to 5) show conserved nucleotide differences from type 1 in the 5'NCR that allow them to be differentiated by restriction endonucleases or by type-specific probes in InnoLipa, many (but not all) of the novel variants from Vietnam and Thailand have sequences in the amplified region used for genotyping that are identical to those of type 1a or 1b. Although mixed infection could lead to the detection of sequences of different genotypes in different amplified regions, this explanation was ruled out by finding continuous clones that contained type 6 group sequences in the core and E1 regions but a 5'NCR sequence similar to that in type 1 in several samples from Vietnam (26). This finding leads to new difficulties when interpreting results from genotyping assays based upon the 5'NCR; samples that appear to be type 1 in the restriction fragment length polymorphism (RFLP) assay (9) or InnoLipa (9) may instead have been variants within the type 6 phylogenetic group.

The core gene encodes a putative nucleocapsid protein, which is the most conserved protein in the HCV genome. Nevertheless, it is considerably more variable than the 5'NCR and contains sufficient sequence information to identify all known subtypes and major genotypes, including those in the type 6 group that are misclassified by using the 5'NCR (15). To investigate the distribution of these novel sequences in Southeast (SE) Asia, we have developed a genotyping assay in which sequences in the core region are amplified and cleaved by restriction enzymes that reliably differentiate type 1 variants from type 6 variants. These results provided information on the distribution of type 6 variants in SE Asia and the frequency with which conventional genotyping assays produced incorrect results in different countries.

## MATERIALS AND METHODS

**Samples.** Plasma samples that had been previously genotyped in the 5'NCR by RFLP (9) from blood donors in Macau, Malaysia, Singapore, Taiwan, and Thailand were used in the current study. Additional samples from Hong Kong ( $n = 66$ ), Malaysia ( $n = 8$ ), and Burma (Myanmar) ( $n = 8$ ) were also obtained and genotyped by the same method prior to analysis in the core region. From the blood donor population in Hong Kong, genotypes 1a (Myanmar) ( $n = 3$ ), 1b ( $n = 40$ ), 2 ( $n = 1$ ), and 6a ( $n = 22$ ) were present, similar to the results of our previous survey (9). From Malaysia, genotypes 1a ( $n = 10$ ), 1b ( $n = 8$ ), and 3 ( $n = 9$ ) were found upon combining the original and new samples. From Burma, types 1b ( $n = 3$ ), 2 ( $n = 2$ ), and 3 ( $n = 3$ ) were identified.

The survey group in the current study was confined to samples of genotype 1 in the 5'NCR RFLP assay that could be amplified in the core region (see below). From Thailand, it was possible to amplify 28 of the 34 type 1 samples in the core region. All of the samples from Burma ( $n = 3$ ), Hong Kong ( $n = 43$ ), and Malaysia ( $n = 18$ ) could be amplified. Lower frequencies of amplification in the core region were obtained from the survey countries surveyed: Macau, 8 of 13; Taiwan, 6 of 53; and Singapore, 2 of 5. Although the core primers are similar in sensitivity to those in the 5'NCR, it is likely that storage conditions influenced the effectiveness of the PCR for the amplification of samples from some countries. A total of 108 samples from all of the countries were used. Previously published sequences from Vietnam and Thailand (1, 6, 26), as well as 25 new core sequences for type 6a (17a), were also included in the analysis.

**RNA extraction and nested PCR amplification.** Virus RNA was extracted directly from 100  $\mu$ l of serum with proteinase K-sarcosyl; this was followed by phenol-chloroform extraction and precipitation in ethanol as previously described (9). RNA was reverse transcribed by using the specific outer antisense core primer 410. This was followed by the first round of PCR with primers 410 and 954 and a second round with primers 951 and 953, described previously (15), to give a 405-bp product between positions -21 and +383; the numbering of nucleotides is from reference 8. Sequences in the 5'NCR between positions -245 and -72 were obtained by using primers 209, 939, 211, and 940 (9). The NS-5 region was amplified by using primers 1204, 1203, 518, 517, and 123, as described by Mellor et al. (15).

**RFLP analysis.** Cleavage reactions were performed on the reverse transcription-PCR 5'NCR product as described previously (9). The restriction enzymes

A)

a1	94	101	38	24	148
a2	94	101	62		148
a3	94	139	24		148
a4	94	139			172
a5	94	139	24	64	84
a6		233	24		148
a7	94	52	87	24	148
a8	94	52	87		172
a9	94	101	38		172
a10	94	163			148
a11	94	51		260	
a12	94			311	
a13	94	101			210

B)

s1	94	163		148
s2		257		148
s3	94	52	259	
s4	94	52	111	148
s5	94		311	

FIG. 1. Restriction endonuclease cleavage patterns of amplified core sequences using enzymes *Ava*I (A) and *Sma*I (B) for published variants and samples analyzed in the current study. The numbers are the sizes (in base pairs) of the DNA fragments produced.

*Hpa*II and *Dde*I identified nucleotide sequence differences in the 5'NCR between type 1 variants and some type 6 group variants (see Results). Restriction endonuclease cleavage of amplified core sequences was carried out by using the restriction endonucleases *Ava*I and *Sma*I. DNA was electrophoresed through a 4% Metaphor agarose gel in 1 $\times$  TBE (Tris-borate-EDTA) containing 0.5  $\mu$ g of ethidium bromide per ml. The combined results of the core and 5'NCR RFLP analyses were used to predict the genotype of HCV within the samples.

**Direct sequencing of PCR products.** Samples showing restriction patterns different from those predicted from published sequence data were reamplified by using biotinylated primers. In order to obtain single-stranded DNA for sequencing, the biotinylated products were bound onto paramagnetic streptavidin-coated beads (Dynabeads M280; Dynal) as previously described (12). Dideoxy-termination sequencing reactions were performed by using the Sequenase sequencing kit from United States Biochemical Corp. according to the manufacturer's instructions, except that reactions were carried out in 10% dimethyl sulfoxide and the template DNA was heat denatured before primer annealing. Sequences were read manually from autoradiograms and analyzed by using standard sequence software.

**Phylogenetic analysis.** Phylogenetic analysis was carried out by using the program NEIGHBOR in the PHYLIP package, as previously described (20).

**Nucleotide sequence accession numbers.** New sequences obtained in this study have been submitted to GenBank and have been assigned accession numbers L49473 to L49485, respectively.

## RESULTS AND DISCUSSION

**Development of core RFLP assay.** Published nucleotide sequences in the core region of types 1 and 6 and novel variants that grouped with type 6 upon phylogenetic analysis were analyzed with DNANALYZ software to identify restriction enzymes that recognized sequence polymorphism between HCV genotypes. Cleavage with *Ava*I and *Sma*I consistently showed distinct predicted restriction patterns for type 1 and non-type 1 sequences (Fig. 1; Table 1). From the total of 13 different restriction patterns obtained using *Ava*I (a1 to a13 [Fig. 1A]) and the five patterns obtained using *Sma*I (s1 to s5 [Fig. 1B]), published sequences of the type 6 group produced a total of three combinations of patterns.

TABLE 1. Association of different HCV genotypes with RFLP patterns in the core region

Genotype	No. of samples with restriction pattern <sup>a</sup> :														Total no. of samples
	a1s1	a2s1	a3s1	a4s1	a4s5	a5s1	a6s5	a7s4	a8s3	a9s2	a10s1	a11s5	a12s5	a13s5	
1a	14	2													16
1b			43		5										48
6a													30		30
T6G <sup>b</sup>											2		15	1	18
Total no. of samples	14	2	43		5						2		45	1	112

<sup>a</sup> Combination of restriction patterns as defined in Fig. 1; a1s1 corresponds to pattern 1 with *Ava*I and pattern 1 with *Sma*I.

<sup>b</sup> Variants of HCV within the type 6 clade but excluding the previously classified type 6a genotype.

Among published sequences, the most common restriction pattern for type 1a sequences was a1s1 ( $n = 14$ ), while the majority of type 1b sequences showed the a3s1 pattern ( $n = 43$ ) (Table 1). With the exception of three sequences (VN540, VN787, and VN507), all variants in the type 6 group (including type 6a) could be predicted to produce the a12s5 pattern upon cleavage with *Ava*I and *Sma*I (Table 1). In every case, RFLP patterns distinct from those of type 1 variants were obtained.

**Geographical distribution of surveyed samples.** To investigate the distribution of type 6 group variants in SE Asia, a total of 108 samples from Malaysia, Thailand, Macau, Hong Kong, Singapore, Burma, and Taiwan that had been previously identified as type 1 (9) were retested in the new core RFLP assay (Table 2). A much wider range of restriction patterns than was obtained from our analysis of published sequences was obtained from this survey. Although the vast majority of samples produced RFLP pattern a1s1 or a3s1 and could therefore be identified as type 1a or 1b, the remaining 26 samples produced a total of 11 distinct RFLP patterns with the two enzymes, many of which corresponded to those previously associated with type 6 group variants (Table 1).

Each of the 26 samples was sequenced in the core region to make a definitive identification of the genotype. Of the 26 samples, 18 could be classified as type 1, while the remaining 8 contained sequences in the core region that placed them within the type 6 group (see below). These type 6 group sequences were found in samples from Thailand ( $n = 7$ ) or Burma ( $n = 1$ ) and showed the restriction pattern a11s5 or a12s5 (Table 2).

**Phylogenetic analysis of novel genotypes.** To demonstrate the relationship between the eight type 6 group variants with those previously found in Vietnam and Thailand, we carried out phylogenetic analysis of sequences from the core and NS-5 regions of these variants (Fig. 2 and 3). All eight core se-

quences grouped with the type 6 clade and showed various relationships with variants previously found in Vietnam and Thailand (13, 15) (Fig. 2). Three samples from Thailand (EUTH5230, EUTH7, and EUTH13) grouped closely with variants previously found in Thailand and provisionally classified as NGII (15). The sample from Burma was closest to but distinct from variants in Thailand referred to as NGI, while the remaining four samples (EUTH1, EUTH21, EUTH22, and EUTH39) showed a distant relationship to variants from Vietnam described as type 9a (13).

To investigate further the relationship of the new variants to those previously found in SE Asia, we compared sequences in the more variable NS-5 region (Fig. 3). Each new sequence clustered within the type 6 group, although the addition of the new sequences from Thailand increased still further the diversity within the group as a whole. Sequence relationships found in the core region among members of the type 6 group were closely reproduced in the NS-5 region. For example, the NS-5 sequences of EUTH1 and EUTH22 were distinct from all previously analyzed variants in the type 6 group but were closest to VN004 and VN085 from Vietnam, which were described as type 9a (13). Similarly, the close relationship between EUTH5230, EUTH7, and EUTH13 with NGII in the core region was also found in NS-5. The NS-5 sequence of the sample from Burma (EUBUR1) was found to be the same genotype as variants B4-92 and PC found in northern Thailand (1).

**Identification of type 6 group sequences in the 5'NCR.** All published type 6a sequences in the 5'NCR are distinct from type 1 sequences, because they contain characteristic insertions (CA at position -146 and C at position -138) which are generally not present in other genotypes (22). However, other variants in the type 6 group show a range of different 5'NCR sequences, none of which contain the CA insertion at position -146 and which infrequently show the C insertion at position

TABLE 2. Distribution of different core RFLP patterns in the places surveyed

Place	No. of samples with restriction pattern <sup>a</sup> :														Total no. of samples
	a1s1	a2s1	a3s1	a4s1	a4s5	a5s1	a6s5	a7s4	a8s3	a9s2	a10s1	a11s5	a12s5	a13s5	
Burma			2										1		3
Thailand	9		9	1	1		1					4	3		28
Malaysia	9	1	5					2	1						18
Singapore			1		1										2
Hong Kong	3		38	1			1								43
Macau	1		2	1		1				1	2				8
Taiwan			3			1					2				6
Total no. of samples	22	1	60	3	2	2	2	2	1	1	4	4	4	0	108

<sup>a</sup> Combination of restriction patterns as defined in Fig. 1; a1s1 corresponds to pattern 1 with *Ava*I and pattern 1 with *Sma*I.

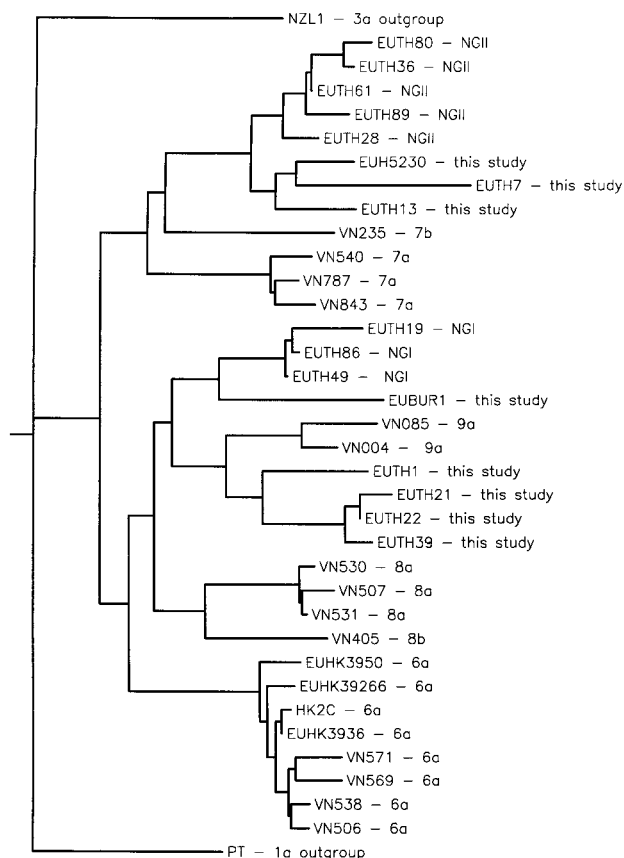


FIG. 2. Phylogenetic analysis of core sequences from the type 6 clade including those of eight variants identified in the current group. The phylogenetic tree was rooted by using sequences from HCV-PT (type 1a [8]) and NZLI (type 3a [18]) as outgroups. Previously published provisional genotype assignments are indicated.

–138. Furthermore, few nucleotide differences are found between these variants and type 1. In the extreme case, for VN540, VN843, VN235, VN507, VN530, and VN531, the 5'NCR sequences were identical to those of type 1b and the 5'NCR sequence of VN085 was identical to that of type 1a (Fig. 4). On the other hand, most of the variants from Thailand showed 5'NCR sequences that differed at several sites from type 1 sequences; in the case of sequences previously described as NGI (15), a conserved pair of covariant changes in the proposed stem-loop region 3 (3) that could be identified by cleavage with *HpaII* (which recognizes the CCGG sequences present in type 1 sequences but which is disrupted by the substitution of A for C at position –155 in NGI sequences) was found. Cleavage of 5'NCR sequences using the combination of *DdeI* and *HpaII* from NGI produced pattern B (Fig. 5), clearly distinguishable from pattern A found in all type 1 5'NCR sequences analyzed to date. A different covariant change in the same stem-loop structure was found among the sequences described as NGII. In this case, cleavage with *DdeI* identified the polymorphism at position –127 (G for A in NGII) and produced pattern C with the *DdeI-HpaII* enzyme combination (Fig. 5). Although pattern D (produced by type 6a 5'NCR sequences [22]) is different from pattern A for type 1, the 3-base size difference of the large fragment (127 and 130 bp) would not necessarily be identifiable by agarose gel electrophoresis of the cleaved DNA (9).

With this limited ability to identify at least a proportion of

type 6 group variants, we used the *DdeI-HpaII* enzyme combination to cleave amplified 5'NCR sequences from the survey population. All 100 sequences that were classified as type 1 by RFLP or sequence analysis of the core region produced pattern A. Among the 8 samples identified in this study as belonging to the type 6 group, EUBUR1 produced pattern B, while EUTH5230, EUTH7, and EUTH13 produced pattern C. EUTH1, EUTH21, EUTH22, and EUTH39, which were not closely related to any of the variants within the type 6 group, produced pattern A (and had sequences in the 5'NCR similar or identical to type 1b [data not shown]). From this analysis it can be seen that only half of the type 6 group variants would have been identified even when using new enzymes designed to detect observed sequence differences between NGI and NGII with type 1 sequences.

**HCV genotype distribution in SE Asia.** The identification of only eight samples in the type 6 group of HCV genotypes indicates an extremely limited distribution of these variants in most of the places surveyed. Seven of these variants were found in Thailand, consistent with previous observations that they represent one of the principal variants (with type 3) in blood donor and hepatitis C patient populations (1, 15). Type 6 group variants (including type 6a) are well documented in Vietnam (26), and the finding of a single example of this genotype in Burma hints at a wider westward distribution in SE Asia.

In attempts to review current information from other groups

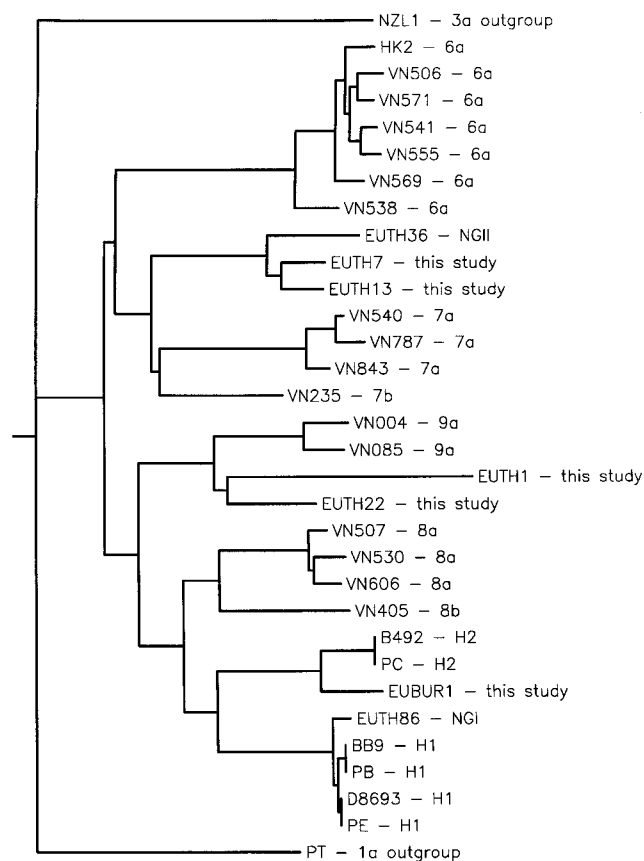
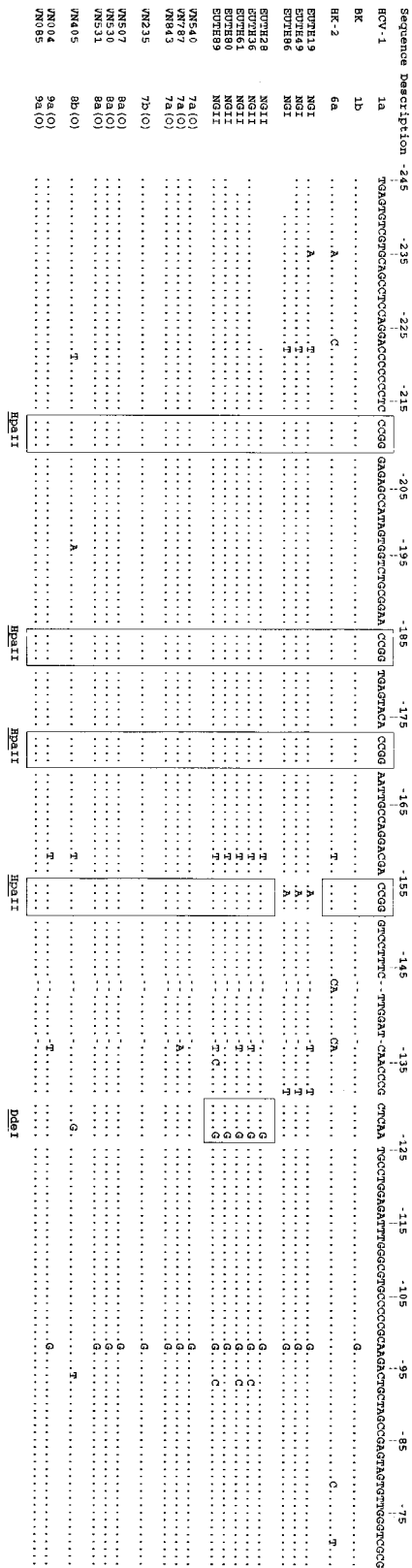


FIG. 3. Phylogenetic analysis of NS-5 sequences from the type 6 clade including those of eight variants identified in the current group. The phylogenetic tree was rooted by using sequences from HCV-PT (type 1a [8]) and NZLI (type 3a [18]) as outgroups. Previously published provisional genotype assignments are indicated.

FIG. 4. Comparisons of sequences in the amplified portion of 5'NCR from type 6 group variants. Nucleotides different from those of HCV-4PT are indicated; nucleotides are numbered as in reference 8. Recognition sequences for the restriction endonucleases *Hpa*II and *Dde*I are in boxes. Sequences are labelled type 7a, 7b, 8a, 8b, 9a, 9b, NGI, and NGII (Description column) according to their original descriptions (15, 26).



A	67	27	12	18	127	
B	67	27	12	145		
C	67	27	12	18	25	102
D	67	27	12	18	130	

FIG. 5. DNA fragments produced upon cleavage of amplified 5'NCR sequences from type 1 and type 6 groups using restriction endonucleases *Dde*I and *Hpa*II. The numbers are the sizes (in base pairs) of the DNA fragments produced.

on the distribution of HCV genotypes in SE Asia, most data obtained from genotyping assays such as the RFLP assays (9), InnoLipa (23), and assays using type-specific primers in the core region (17) cannot be relied upon because of the limited range of HCV variants that can be identified. In the current study, we have shown that type 6 group variants may appear as type 1 in an RFLP assay based upon the 5'NCR, whether or not it is modified by the use of additional restriction enzymes such as *Dde*I and *Hpa*I. In this respect, the survey recently published by Greene et al. (10) is particularly informative, because genotype assignments were made by nucleotide sequence determinations of several regions of the HCV genome. The absence of type 6 group variants from a total of 45 samples collected from Singapore, Indonesia, Philippines, and South Korea is consistent with the finding in this study of a restricted distribution of type 6 group variants in Malaysia, Singapore, Taiwan, Macau, and Hong Kong. However, the survey by Greene et al. (10) failed to detect type 6 group variants in Thailand (3 type 1a, 8 type 1b, and 10 type 3a and 3b variants) and appears inconsistent with their frequent detection in this study and previous studies (1, 15). It is possible that the frequency of infection with type 6 group variants differs between geographical regions within Thailand, or it is possible that there are differences in its distribution in different risk groups for infection or age of subjects.

Figure 6 attempts to summarize current knowledge of genotype distributions of HCV throughout SE Asia and the Far East using information obtained from this study and previously published information (9, 10, 26). The pie charts represent the relative frequencies of genotypes 1 to 6 as determined by sequence comparisons in coding regions (9, 15, 26), RFLP analysis in the 5'NCR (9), and genotyping using type-specific primers in the core region (26). Figure 6 includes a separate representation of type 6 group variants (indicated as NG, distinct from type 6a) in the three countries where it has been detected so far (Vietnam, Thailand, and Burma).

Although the number of samples from some places was small, this combined analysis reveals some clear geographical trends in relative genotype frequencies. Type 1b variants appear ubiquitous throughout the region, and a steady increase in the frequency of type 3 infection in more westward countries is apparent, which is not found in the Far East. Type 6 group variants were found in a single, possibly continuous block that includes Vietnam, Thailand, and Burma but were not detected elsewhere. The distribution of type 6 group variants overlaps in Vietnam with type 6a, with the latter also being prevalent in Hong Kong and Macau.

Little is currently understood concerning the past epidemiology of HCV transmission that is responsible for the current distribution of HCV variants. However, the information obtained in this study will be of future use in the understanding of HCV sequence variation and the ancestry of the currently identified HCV genotypes. In particular, the overlapping ranges of type 6a with other variants in the type 6 clade in the mainland of SE Asia provides an insight into their interrela-

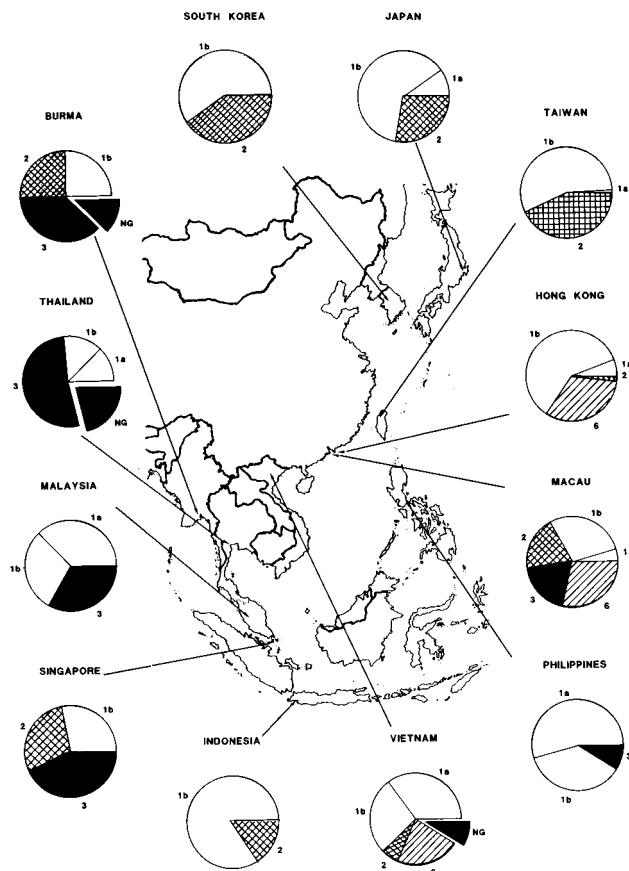


FIG. 6. Map of SE Asia summarizing the HCV genotype distributions identified in the current study and by previously published analyses. Pie charts indicate the proportions of various HCV genotypes in each place surveyed, although the sample numbers available were small in some places, such as Burma. In addition to the current study, data for Indonesia, Philippines, and South Korea were obtained from reference 10, data for Vietnam were from reference 26, and data for Japan were from sequence information in GenBank.

tionship that will be of value in future evolutionary analyses of these sequences.

#### ACKNOWLEDGMENTS

We are very grateful to the following individuals who contributed to the current study by collecting blood samples: Julie Ferguson and Jennifer Daub, Edinburgh and South East Scotland Blood Transfusion Service; Ronald Chi Lang Chan and Hoo Chai of the Macau Blood Transfusion Centre; and W. S. Lin, Taichung Blood Donation Centre, Taichung, Taiwan. We are grateful to J. D. Cash for continued support and encouragement.

The work in Scotland was funded by the Scottish National Blood Transfusion Service and the Medical Research Council (grant number G9020615CA).

#### REFERENCES

- Apichartpiyakul, C., C. Chittivudikarn, H. Miyajima, M. Homma, and H. Hotta. 1994. Analysis of hepatitis C virus isolates among healthy blood donors and drug addicts in Chiang Mai, Thailand. *J. Clin. Microbiol.* **32**: 2276–2279.
- Bhattacharjee, V., L. E. Prescott, I. Pike, B. Rodgers, H. Bell, A. R. Elzayadi, M. C. Kew, J. Conradie, C. K. Lin, H. Marsden, A. A. Saeed, D. Parker, P. L. Yap, and P. Simmonds. 1995. Use of NS-4 peptides to identify type-specific antibody to hepatitis C virus genotypes 1, 2, 3, 4, 5 and 6. *J. Gen. Virol.* **76**: 1737–1748.

- Brown, E. A., H. Zhang, H.-L. Ping, and S. M. Lemon. 1992. Secondary structure of the 5' nontranslated region of hepatitis C virus and pestivirus genomic RNAs. *Nucleic Acids Res.* **20**:5041–5045.
- Bukh, J. 1995. Genetic heterogeneity of hepatitis C virus: quasispecies and genotypes. *Semin. Liver Dis.* **15**:41–63.
- Bukh, J., R. H. Purcell, and R. H. Miller. 1993. At least 12 genotypes of hepatitis C virus predicted by sequence analysis of the putative E1 gene of isolates collected worldwide. *Proc. Natl. Acad. Sci. USA* **90**:8234–8238.
- Bukh, J., R. H. Purcell, and R. H. Miller. 1994. Sequence analysis of the core gene of 14 hepatitis C virus genotypes. *Proc. Natl. Acad. Sci. USA* **91**: 8239–8243.
- Choo, Q. L., G. Kuo, A. J. Weiner, L. R. Overby, D. W. Bradley, and M. Houghton. 1989. Isolation of a cDNA derived from a blood-borne non-A, non-B hepatitis genome. *Science* **244**:359–362.
- Choo, Q. L., K. H. Richman, J. H. Han, K. Berger, C. Lee, C. Dong, C. Gallegos, D. Coit, R. Medina-Selby, P. J. Barr, A. J. Weiner, D. W. Bradley, G. Kuo, and M. Houghton. 1991. Genetic organization and diversity of the hepatitis C virus. *Proc. Natl. Acad. Sci. USA* **88**:2451–2455.
- Davidson, F., P. Simmonds, J. C. Ferguson, L. M. Jarvis, B. C. Dow, E. A. C. Follett, C. R. G. Seed, T. Krusius, C. Lin, G. A. Medgyesi, H. Kiyokawa, G. Olim, G. Duraisamy, T. Cuypers, A. A. Saeed, D. Teo, J. Conradie, M. C. Kew, M. Lin, C. Nuchprayoon, O. K. Ndimbie, and P. L. Yap. 1995. Survey of major genotypes and subtypes of hepatitis C virus using RFLP of sequences amplified from the 5' non-coding region. *J. Gen. Virol.* **76**:1197–1204.
- Greene, W. K., M. K. Cheong, V. Ng, and K. W. Yap. 1995. Prevalence of hepatitis C sequence variants in South-East Asia. *J. Gen. Virol.* **76**:211–215.
- Holmes, E. C., P. Simmonds, T. A. Cha, S. W. Chan, F. McOmish, B. Irvine, E. Beall, P. L. Yap, J. Kolberg, and M. S. Urdea. 1993. Derivation of a rational nomenclature for hepatitis C virus by phylogenetic analysis of the NS-5 region, p. 57–62. *In* K. Nishioka, H. Suzuki, S. Mishihiro, and T. Oda (ed.), *Viral hepatitis and liver disease*. Springer-Verlag, Tokyo.
- Kleter, G. E. M., L. J. Van Doorn, J. T. Brouwer, S. W. Schalm, R. A. Heijink, and W. G. V. Quint. 1994. Sequence analysis of the 5' untranslated region in isolates of at least four genotypes of hepatitis C virus in The Netherlands. *J. Clin. Microbiol.* **33**:306–310.
- Luengrojnanakul, P., K. Varesangthip, T. Chainuvati, K. Murata, F. Tsuda, H. Tokita, H. Okamoto, Y. Miyakawa, and M. Mayumi. 1994. Hepatitis C virus infection in patients with chronic liver disease or chronic renal failure and blood donors in Thailand. *J. Med. Virol.* **44**:287–292.
- Machida, A., H. Ohnuma, F. Tsuda, E. Munekata, T. Tanaka, Y. Akahane, H. Okamoto, and S. Mishihiro. 1992. Two distinct subtypes of hepatitis C virus defined by antibodies directed to the putative core protein. *Hepatology* **16**: 886–891.
- Mellor, J., E. C. Holmes, L. M. Jarvis, P. L. Yap, P. Simmonds, and International HCV Collaborative Study Group. 1995. Investigation of the pattern of hepatitis C virus sequence diversity in different geographical regions: implications for virus classification. *J. Gen. Virol.* **76**:2493–2507.
- Okamoto, H., K. Kurai, S. Okada, K. Yamamoto, H. Lizuka, T. Tanaka, S. Fukuda, F. Tsuda, and S. Mishihiro. 1992. Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct genotypes. *Virology* **188**:331–341.
- Okamoto, H., H. Tokita, M. Sakamoto, M. Horikita, M. Kojima, H. Iizuka, and S. Mishihiro. 1993. Characterization of the genomic sequence of type V (or 3a) hepatitis C virus isolates and PCR primers for specific detection. *J. Gen. Virol.* **74**:2385–2390.
- Prescott, L. E. Unpublished data.
- Sakamoto, M., Y. Akahane, F. Tsuda, T. Tanaka, D. G. Woodfield, and H. Okamoto. 1994. Entire nucleotide sequence and characterization of a hepatitis C virus of genotype v/3a. *J. Gen. Virol.* **75**:1761–1768.
- Simmonds, P. 1995. Variability of hepatitis C virus. *Hepatology* **21**:570–583.
- Simmonds, P., E. C. Holmes, T. A. Cha, S.-W. Chan, F. McOmish, B. Irvine, E. Beall, P. L. Yap, J. Kolberg, and M. S. Urdea. 1993. Classification of hepatitis C virus into six major genotypes and a series of subtypes by phylogenetic analysis of the NS-5 region. *J. Gen. Virol.* **74**:2391–2399.
- Simmonds, P., K. A. Rose, S. Graham, S.-W. Chan, F. McOmish, B. C. Dow, E. A. C. Follett, P. L. Yap, and H. Marsden. 1993. Mapping of serotype-specific, immunodominant epitopes in the NS-4 region of hepatitis C virus (HCV): use of type-specific peptides to serologically differentiate infections with HCV types 1, 2, and 3. *J. Clin. Microbiol.* **31**:1493–1503.
- Smith, D. B., J. Mellor, L. M. Jarvis, F. Davidson, J. Kolberg, M. Urdea, P. L. Yap, P. Simmonds, J. D. Conradie, A. G. S. Neill, G. M. Dusheiko, M. C. Kew, R. Crookes, A. Koshy, C. K. Lin, C. Lai, I. M. Murraylyon, A. Elgunaid, A. A. Gunaid, T. Yemen, S. Yemen, D. Mutimer, M. Ahmed, C. Nuchprayoon, S. Tanprasert, F. E. Preston, M. Makris, A. Chuansumrit, C. Mahasandana, D. Pritchard, E. Riley, B. M. Greenwood, A. A. Saeed, A. M. Alrasheed, M. G. Saleh, I. McFarlane, C. Tibbs, R. Williams, J. Power, E. Lawlor, and H. Kiyokawa. 1995. Variation of the hepatitis C virus 5' non-coding region: implications for secondary structure, virus detection and typing. *J. Gen. Virol.* **76**:1749–1761.
- Stuyver, L., R. Rossau, A. Wyseur, M. Duhamel, B. Vanderborght, H. Van

- Heuverswyn, and G. Maertens.** 1993. Typing of hepatitis C virus isolates and characterization of new subtypes using a line probe assay. *J. Gen. Virol.* **74**: 1093–1102.
24. **Stuyver, L., W. Vanarnhem, A. Wyseur, F. Hernandez, E. Delaporte, and G. Maertens.** 1994. Classification of hepatitis C viruses based on phylogenetic analysis of the envelope 1 and nonstructural 5b regions and identification of five additional subtypes. *Proc. Natl. Acad. Sci. USA* **91**:10134–10138.
25. **Tanaka, T., K. Tsukiyamakohara, K. Yamaguchi, S. Yagi, S. Tanaka, A. Hasegawa, Y. Ohta, N. Hattori, and M. Kohara.** 1994. Significance of specific antibody assay for genotyping of hepatitis C virus. *Hepatology* **19**: 1347–1353.
26. **Tokita, H., H. Okamoto, F. Tsuda, P. Song, S. Nakata, T. Chosa, H. Iizuka, S. Mishiro, Y. Miyakawa, and M. Mayumi.** 1994. Hepatitis C virus variants from Vietnam are classifiable into the seventh, eighth, and ninth major genetic groups. *Proc. Natl. Acad. Sci. USA* **91**:11022–11026.
27. **Van Doorn, L. J.** 1994. Detection and characterisation of the hepatitis C virus. Ph.D. dissertation. Erasmus Universiteit Rotterdam, Rotterdam, The Netherlands.